Problem 1

a) According to Bayes rule:

$$p(\varphi_{ij}|r_{ij},u_i,v_j) = \frac{p(r_{ij}|\varphi_{ij})\ p(\varphi_{ij}|u_i,v_j)}{\int p(r_{ij}|\varphi_{ij})\ p(\varphi_{ij}|u_i,v_j)d\varphi_{ij}}$$

According to the model $r_{i,j} = \mathrm{sign}(\varphi_{i,j})$:

$$p(r_{ij} = 1|\varphi_{ij}) = 1(\varphi_{ij} > 0)$$

$$p(r_{i,j} = -1|\varphi_{ij}) = 1(\varphi_{ij} \le 0)$$

Therefore, when $r_{i,j} = 1$, $p(\varphi_{ij}|r_{ij}, u_i, v_j) = TN_1(u_i^T v_j, \sigma^2)$ and when $r_{ij} = -1$, $p(\varphi_{ij}|r_{ij}, u_i, v_j) = TN_{-1}(u_i^T v_j, \sigma^2)$

Then, we can derive $q(\varphi)$ by :

$$\begin{aligned} q(\varphi) &= \prod_{ij\in\Omega} q(\varphi_{ij}) \\ &= \prod_{ij\in\Omega} p(\varphi_{ij}|r_{ij},u_i,v_j) \\ &= \prod_{ij\in\Omega} TN_{r_{ij}}(u_i^T v_j, \sigma^2) \end{aligned}$$

b) According to EM algorithm:

$$\mathcal{L}(U,V) = \int q(\varphi)\ln\frac{p(R,U,V,\varphi)}{q(\varphi)}d\varphi$$

According to Bayes rule:

$$\mathcal{L}(U,V) = \int q(\varphi)(lnp(U,V) + lnp(R|\varphi) + \ln p(\varphi|U,V))d\varphi - E_q[lnq(\varphi)]$$

Since $q(\varphi_{i,j}) = TN_{r_{i,j}}(u_i^T v_j, \sigma^2)$ is only defined on values of φ_{ij} when $\mathrm{sign}(\varphi_{ij}) = r_{ij}$, $\mathrm{lnp}(R|\varphi) = 0$. And $E_q[lnq(\varphi)]$ is a constant . Then we can get:

$$\mathcal{L}(U,V) = \ln p(U,V) + E_q[\ln p(\varphi|U,V)] - E_q[lnq(\varphi)]$$

Because of the independence of each u_i, v_j:

$$E_q[\ln p(\varphi|U,V)] = \sum_{i,j\in\Omega} E_q[\ln p(\varphi_{ij}|u_i,v_j)]$$

$$\ln p(U,V) = -\frac{1}{2c}\left(\sum_{i=1}^{M} u_i^T u_i + \sum_{j=1}^{N} v_j^T v_j\right) - \frac{(M+N)d}{2}\ln 2\pi c$$

Then we can get:

$$\mathcal{L}(U,V) = -\frac{1}{2c}\left(\sum_{i=1}^{M} u_i^T u_i + \sum_{j=1}^{N} v_j^T v_j\right) + \sum_{i,j\in\Omega} E_q\left[-\frac{(\varphi_{ij}-u_i^T v_j)^2}{2\sigma^2}\right] + const$$

$$= -\frac{1}{2c}\left(\sum_{i=1}^{M} u_i^T u_i + \sum_{j=1}^{N} v_j^T v_j\right) + \sum_{i,j\in\Omega}\frac{1}{2\sigma^2}\left(2u_i^T v_j E_q\left[\varphi_{ij}\right] - (u_i^T v_j)^2\right) + const$$

c)

Solving for $\nabla_{U,V}\ \mathcal{L}(U, V) = 0$, we can find that one solution is:

$$u_i = (\frac{1}{c}I + \sum_{(i,j)\in\Omega} v_j v_j^T/\sigma^2)^{-1} \sum_{(i,j)\in\Omega} \mathrm{E_q}[\varphi_{ij}] v_j/\sigma^2$$

$$v_j = (\frac{1}{c}I + \sum_{(i,j)\in\Omega} u_i u_i^T/\sigma^2)^{-1} \sum_{(i,j)\in\Omega} \mathrm{E_q}[\varphi_{ij}] u_i/\sigma^2$$

d)

1. Initializing U_0, V_0 to vectors generated from normal distribution
2. For interation t = 1,2, ..., T
   (a)E-step: Calculate $\mathrm{E} = \{E_q[\varphi_{i,j}]\}$, where

$$E_q[\varphi_{ij}] = \begin{cases} u_{i(t-1)}^T v_{j(t-1)} + \sigma \times \dfrac{\Phi'(-\frac{u_{i(t-1)}^T v_{j(t-1)}}{\sigma})}{1-\Phi(-\frac{u_{i(t-1)}^T v_{j(t-1)}}{\sigma})} & \text{if } r_{ij} = 1 \\ u_{i(t-1)}^T v_{j(t-1)} + \sigma \times \dfrac{-\Phi'(-\frac{u_{i(t-1)}^T v_{j(t-1)}}{\sigma})}{\Phi(-\frac{u_{i(t-1)}^T v_{j(t-1)}}{\sigma})} & \text{if } r_{ij} = -1 \end{cases}$$

   (b)M-step: Update vector U, V using the expectations above in the following equations

$$u_{it} = (\frac{1}{c}I + \sum_{(i,j)\in\Omega} v_{j(t-1)} v_{j(t-1)}^T/\sigma^2)^{-1} \sum_{(i,j)\in\Omega} \mathrm{E_{qt}}[\varphi_{ij}] v_{j(t-1)}/\sigma^2$$

$$v_{jt} = (\frac{1}{c}I + \sum_{(i,j)\in\Omega} u_{it} u_{it}^T/\sigma^2)^{-1} \sum_{(i,j)\in\Omega} \mathrm{E_{qt}}[\varphi_{ij}] u_{it}/\sigma^2$$

   (c)Calculate ln p(U,V,R) using the equation

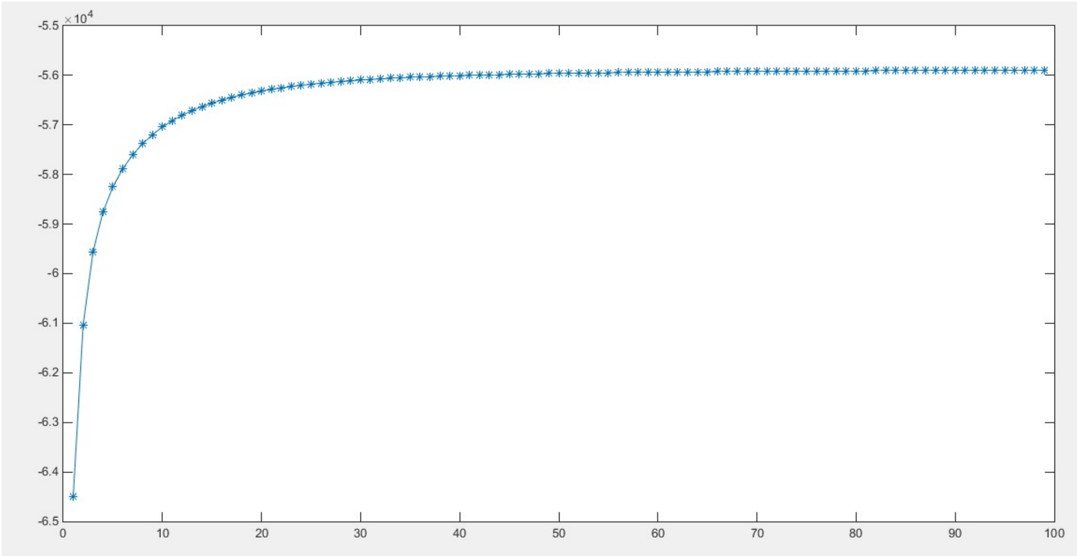
$$\ln p(U,V,R) = -\frac{(M+N)d}{2}\ln 2\pi c - \frac{1}{2c}(\sum_{i=1}^{M} u_i^T u_i + \sum_{j=1}^{N} v_j^T v_j) + \sum_{i,j\in\Omega} \ln(u(-r_{ij}) + r_{ij}\,\Phi(\frac{u_i^T v_j}{\sigma}))$$

   u(x) is the step function

Problem 2

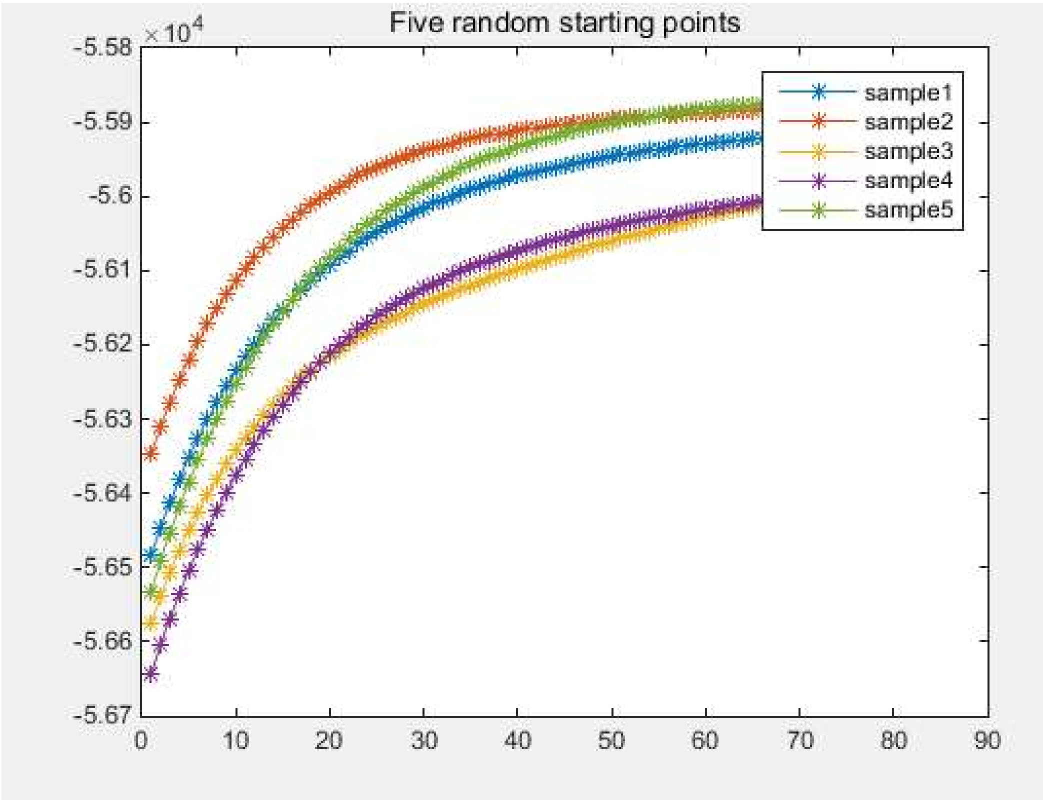
a)

Value of ln p(U,V,R) for iterations 2 through 100



b)

Value of ln p(U,V,R) for iterations 20 through 100 with 5 random starting points



c)

confusion matrix:

	predicted_like	predicted_dislike
like	2145	589
dislike	824	1442